

УДК 81-13:811.11

РОЛЬ КОМП'ЮТЕРНОЇ ЛІНГВІСТИКИ У НАВЧАННІ ТА ВДОСКОНАЛЕННІ НАВИКІВ ПЕРЕКЛАДУ

Греков В.О.

Чорноморський державний університет імені Петра Могили

Дана стаття містить аналіз ролі такої дисципліни, як комп'ютерна лінгвістика, у навчанні та роботі професійного перекладача. Подано графічні дані, що наочно ілюструють матеріал, причини викладання даної дисципліни у ВНЗ для спеціальності «Переклад» та на факультетах лінгвістичного спрямування, а також проаналізовано на конкретних прикладах, як комп'ютерна лінгвістика становлять інтерес для початківця і професійного перекладача, і чим саме. Стаття містить тлумачення базових понять та термінів, а також приклади і історію виникнення принципів комп'ютерного перекладу, із поясненням необхідності їх вивчення саме для перекладачів. Досліджено механізм залучення лінгвістичних даних у процес перекладу. **Ключові слова:** мовознавство, комп'ютерна лінгвістика, переклад, професійний перекладач, алгоритм, корпус мови, розпізнавання мовлення, синтез мовлення.

Комп'ютерна лінгвістика є дисципліною, кобов'язковою для вивчення на філологічних факультетах, причому саме для перекладачів вона становить значно більший інтерес, позаяк вони уможливають здійснення перекладів швидше та з більшою якістю, тому перекладач має бути обізнаним у інформаційно-технологічних лінгвістичних аспектах не однієї мови, як філолог, а одразу двох, що суттєво збільшує навантаження. Проблемою даної статті є висвітлення того, чому комп'ютерна лінгвістика становить особливий інтерес для початківців, професійних перекладачів та студентів спеціальності «Переклад», оскільки це допоможе студентові більш ефективно структурувати процес навчання, а фахівцеві – доповнити проміжки у наявних знаннях.

Незважаючи на те, що комп'ютерна лінгвістика як мовознавча дисципліна виникла і сформувалася лише у другій половині двадцятого сторіччя, власне темі гендерної лінгвістики було присвячено велику кількість публікацій, зокрема радянських [1, 2, 5, 10], українських [7, 9], зарубіжних [4] та російських [3, 4, 6, 8, 11, 12]. Проте жодне з цих джерел не розглядає питання ролі, що її відіграє комп'ютерна лінгвістика для перекладачів, ані містить відомості, корисні для перекладачів, що видаються найбільш важливими для залучення у практичну перекладацьку діяльність. **Це зумовлює актуальність даної статті.**

Отже, метою даної статті є:

- Прояснити роль дисципліни «комп'ютерна лінгвістика» в навчанні професійних перекладачів;
- Назвати причини, з яких знання з комп'ютерної лінгвістики є необхідним інструментом у роботі перекладача як міжкультурного комунікатору;
- Довести необхідність вивчення перекладачем певних галузей комп'ютерної лінгвістики та практичного застосування засвоєного в процесі вивчення даної дисципліни теоретичного матеріалу.

Виклад основного матеріалу. Перші серйозні відомі роботи у галузі комп'ютерної лінгвістики почали з'являтися ще у п'ятдесяті роки ХХ століттях [10, 2]. Зумовлені вони були розвитком технічного прогресу та поступовим залученням комп'ютерів – які на той момент ще називали ЕОМ

(електронно-обчислювальна машина) – до усіх галузей людської культури. Цілком зрозуміло, що лінгвісти також стали задумуватися, як використати незвіданий тоді потенціал комп'ютерів для власних цілей. Потужності ЕОМ стосовно обробки інформації у цифровому вигляді викликали подив уже тоді, тож було створено спеціальний підрозділ лінгвістики – комп'ютерна лінгвістика, однією із сфер дослідження якої був і машинний, або автоматизований, переклад. Саме тоді, ще у далекі 50-ті роки, з'явилися зародки ери, яка у майбутньому залишить перекладачів без роботи і зведе їх роль у кращому випадку до операторів машини, яка буде використовувати алгоритми для перекладу кращого, ніж людина змогла б виконати. Але для того, щоб краще зрозуміти повноту значення комп'ютерної лінгвістики, слід прояснити сутність цієї дисципліни, механізми уможливлення її появи і перспективи розвитку, а також значення її як для філологів, так і власне для перекладачів.

Отже, технологічним відкриттям, яке знаменувало прорив у вигляді і принципах функціонування ЕОМ, став винахід транзисторів на заміну застарілим електронним лампам як компоненту схем комп'ютерів. Транзистор – це радіоелектронний компонент, зроблений із напівпровідникового матеріалу, що використовується для підсилення, генерації, комутації та перетворення електричних сигналів. Наразі саме транзистори лежать в основі електронних пристроїв, мікросхем і більшості техніки, що містить електронні компоненти. За свій винахід, представлений 23 грудня 1947 року, У. Шоклі, Д. Бардін та У. Браттейн отримали 1956 року Нобелівську премію. Транзистори не тільки підвищили швидкість обміну електронними імпульсами з комп'ютером, але й мініатюризували їх, зробивши значно компактнішими, як ми знаємо їх зараз. Транзистори зазнали другого етапу еволюції у 1990-ті роки, коли конструкцію було вдосконалено до «біполярного транзистора», меншими за розміром і значно потужнішими. Сьогодні ведуться активні розробки по упровадженню напівпровідникового матеріалу для транзисторів, що міг би ще пришвидшити їх роботу, пропонують використовувати принципово новий синтетичний матеріал графен. Наука та прогрес у галузі фізичного боку

комп'ютера не стоїть на місці, обіцяючи нам нові відкриття і можливості.

Після винаходу і напроцуд швидкої появи комп'ютерів нового покоління наступним серйозним кроком був винахід другого компоненту схеми, не-фізичного. Це інтелектуальна схема, що власне і стоїть за усіма здобутками комп'ютерної лінгвістики. Взагалі будь-яка взаємодія людини і комп'ютера можлива за умов наявності фізичного інтерфейсу (hardware) та програмного інтерфейсу (software). І якщо за фізичний бік революції у комп'ютерній науці відповідали транзистори, саме інтелектуальна схема відповідає за програмний бік взаємодії.

Ще у 1950-ті роки було виведено концепцію, згідно з якою структура інтелектуальної комп'ютерної системи включає в себе три основні блоки (рис. 1).



Рис. 1.

База знань представлена у вигляді фізичного накопичувача, на якому міститься у цифровому вигляді, придатному для взаємодії, інформація, якою буде оперувати програмне забезпечення.

Логічний блок, або вирішувач, має за основну функцію пошук виводу, відповіді на вхідний запит, що потрапляє до системи. Якщо в базі знань (базі даних) міститься інформація, що відповідає поставленому питанню або запиту, сформованому природною мовою, система комунікації перетворює природну мову на її внутрішній цифровий аналог, з яким і працює логічний блок, розшифровуючи його як запит на доступ до бази знань. У випадку, якщо логічний блок не знаходить прямої відповіді на поставлений запит, він за вбудованою функцією влаштовує пошук непрямої інформації, яка так чи інакше стосується запиту.

Інтелектуальний інтерфейс забезпечує взаємодію користувача з комп'ютером та коректне функціонування компонентів інтелектуальної комп'ютерної схеми. Частиною, а нині основою і ядром інтелектуального інтерфейсу багатьох програм є штучний інтелект.

Штучним інтелектом називають властивості комп'ютерних інтелектуальних систем вдаватися до творчих функцій, що традиційно вважаються притаманними лише людині. Сучасні дослідники поділяють ШІ на сильний та слабкий. Сильний представляє собою багато ШІ, інтегрованих в об'єднану систему, достатньо потужну, щоб вирішувати навіть загальнолюдські проблеми. Слабкий – це вузькоспеціалізований ШІ,

інтегрований з метою допомагати людині в певних конкретних галузях. У комп'ютерній лінгвістиці використовується концепт слабого ШІ.

Серед світових лідерів у розробці систем штучного інтелекту можна назвати дослідницькі центри в США – Масачусетський технологічний інститут (Massachusetts Institute of Technology, MIT), Дослідницький інститут машинного інтелекту (Machine Intelligence Research Institute, MIRI), – Японії – Національний інститут сучасної промислової науки і технології (AIST), а також Німеччині – Німецький дослідницький центр з питань штучного інтелекту (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI).

Найкориснішим серед властивостей штучного інтелекту є те, що він здатен до навчання і самонавчання. Над цим працюють відразу кілька спрямувань на перехресті інформаційних технологій та лінгвістики. Схема машинного навчання – це штучний інтелект, первинною спрямованістю якого є не пряме вирішення поставленого завдання, а навчання в процесі його застосування того, як вирішувати багато схожих завдань. Розрізняють два методи навчання ШІ: навчання за прецедентами і дедуктивне навчання.

Навчання за прецедентами ґрунтується на аналізі комп'ютером кінцевої сукупності факторів – прецедентів – на основі навчальної вибірки, на основі яких ШІ має побудувати алгоритм, що видає достатньо правильну відповідь на кожен із об'єктів. Для вимірювання правильності відповідей в ШІ вбудовано також функціонал якості.

Дедуктивне навчання є прерогативою так званих експертних систем, і є наступним етапом – коли машина, що вже навчена на прецедентах, виконує роль експерта, аналізує наявні дані на основі запиту і видає експертний вирок.

Отже, прояснивши те, розробка чого саме уможливила появу комп'ютерної лінгвістики, переходимо до того, в яких галузях лінгвістики комп'ютер допомагає людині, і яким чином (рис. 2).



Рис. 2.

Надана вище схема у спрощеному вигляді допомагає розкрити спектр функцій, за допомогою яких комп'ютер допомагає сучасним лінгвістам.

Функція розпізнавання й оцифрування у широкому сенсі розуміння – це здатність комп'ютера за допомогою інтерфейсу, фізичних сенсорних пристроїв або програмних аналітичних засобів розпізнавати інформацію, наприклад:

– Перетворення мовленнєвого сигналу на цифрову інформацію. Перший пристрій для розпізнавання мовлення з'явився 1952 року, він міг розпізнавати цифри, що їх вимовляв користувач. Комерційно використані програми з розпіз-

навання мовлення з'явилися у 1990-х, частково вони були покликані полегшити користування комп'ютером для тих, хто не міг користуватися клавіатурою як засобом вводу інформації внаслідок перенесених травм або обмежених можливостей. Серед цих програм можна назвати Dragon NaturallySpeaking, VoiceNavigator, Microsoft Voice Command та інші. Механізм роботи такий: обробка мовлення починається з оцінки якості мовленнєвого сигналу, визначається рівень поміх та спотворень. Результат оцінки надається до модулю акустичної адаптації, який керує модулем розрахунку параметрів, необхідних для розпізнавання. Далі у сигналі виділяються ділянки, що містять мовлення, і оцінюються його параметри, виділяються фонетичні та просодичні характеристики, проводиться синтаксичний, семантичний і прагматичний аналіз. Нарешті, проаналізована інформація поступає до декодера, що зіставляє вхідний мовленнєвий потік із інформацією, наявною в акустичній та мовній базі даних, і видає найбільш імовірну послідовність слів, що і стане кінцевим результатом (рис. 3).



Рис. 3.

– Переведення рукописного тексту природною мовою в цифрову інформацію;

– Розпізнавання тексту природною мовою на зображення різного цифрового формату;

– Розпізнавання жестів, проведених у режимі реального часу на камеру (сенсорний пристрій) або попередньо записаних у відеоформаті.

За різних варіантів розпізнавання два останні етапи, зображені у схемі для розпізнавання мовлення, залишаються незмінними, тоді як два перших варіюватимуться в залежності від типу інформації, яку треба розпізнати.

Наступною функцією є статистика та каталогізація. Перед комп'ютерними лінгвістами постає питання, як формалізувати інформацію, що наразі наявна у розпізаному електронному вигляді. Цим займається корпусна лінгвістика – розділ мовознавства, що займається розробкою, створенням і використанням текстових корпусів, тобто сукупностей текстів, відібраних відповідно до певних принципів. Ці принципи каталогізації закладені у ШІ, що відповідає за статистичну обробку та каталогізацію текстів, що поступають до бази знань. Необхідність корпусів текстових даних пояснюється тим, що це є представленням лінгвістичної інформації в конкретному визначеному контексті; за великого об'єму корпусу маємо також велику наявність інформації для аналізу; нарешті, одноразово створений корпус має багаторазове застосування для вирішення різноманітних лінгвістичних записів і проблем. Мовою інформаційних технологій, корпусна лінгвістика займається створенням та укладанням гарно структурованої бази даних, що стосується мови у її конкретних проявах – усних та письмових актах комунікації. Першим створеним лінгвістичним корпусом був так званий корпус Брауна, створений на початку 1960-тих років у універси-

теті Брауна. Він містив 500 фрагментів, кожен із яких складався з 2000 слів англійської мови. Цей стандарт розміру корпусу слів мови – 1 000 000 – був використаний при створенні аналогічних корпусів у 80-ті роки ХХ століття для російської, німецької та французької мов. Наразі в інтернеті наявний сайт Татоеба, що має на меті на вільній основі додавати та змінювати речення та їх переклади, пов'язані між собою за змістом, тобто об'єднувати корпуси двох мов у єдине перекладацьке середовище. Цей ресурс може стати у пригоді як комп'ютерним лінгвістам, так і професійним перекладачам, оскільки налаштування корпусів пари мов для віднайдення чітких еквівалентів і є однією з основ роботи машинного перекладу. Наразі кількість мов на ресурсі більше 80, а кількість речень перевищила відмітку в 600 000. Наявна також функція безкоштовного завантаження всіх корпусів – цілком або частково.

Аналітика і висновки. Існує загальний напрям розвитку ШІ на засадах математичної лінгвістики, що називається обробкою природної мови (англ. NLP, Natural Language Processing). Ця дисципліна вивчає проблеми комп'ютерного аналізу та синтезу природних мов. У контексті ШІ аналіз означає розуміння мови, а синтез – генерацію на основі висновків якісного тексту. Розуміння природної мови залежить від надзвичайно великої кількості факторів – лінгвістичних, екстралінгвістичних (культурних, соціологічних, історичних), особистості співрозмовника і под. Серед лінгвістичних складнощів, з якими стикається ШІ на етапі розуміння тексту, можна навести такі:

– Розкриття анафор (тобто розуміння того, що мається на увазі за умови застосування займенників). Наприклад, є два речення: «Ми видалили дані співробітників, тому що вони були пошкоджені» і «Ми видалили дані співробітників тому, що вони були скомпрометовані». У першому випадку займенник «вони» стосується слова «дані», у другому – слова «співробітники». Правильне розуміння комп'ютером змісту речення залежить від того, чи знає він, яка характеристика підходить до якого з іменників.

– Вільний порядок слів української мови може призвести до неправильного розуміння фрази: «Порядок визначає процес» – незрозуміло, що визначається чим, чи порядок процесом, або ж навпаки, процес – порядком.

– Наявність у тексті неологізмів типу «есемеснути», «зафрендити», які ШІ має відрізнити від орфографічних та друкарських помилок і розуміти.

– Омоніми типу «ключ\ключ», «зАмок\замОк» і под. становлять проблему, оскільки потребують детального аналізу і розуміння мовленнєвого оточення, тобто контексту, перш ніж використовувати (або перекладати) їх.

Серед наявних систем для аналізу та обробки природного тексту можна назвати AlchemyAPI, Natural Language Toolkit, MontyLingua, General Architecture for Text Engineering (GATE).

Синтез і перекодування. Під синтезом у даному контексті розуміємо генерацію якісного тексту комп'ютером, а під перекодуванням – додатковий етап перекладу даного тексту іншою мовою перед етапом його фінальної генерації. Саме ця функція комп'ютерних інтелектуальних схем із

застосуванням ШІ і подарувала нам програми для машинного та автоматизованого перекладу, на чому детальніше зосередимося нижче. Повний синтез мовлення за усіма правилами (або синтез за друкованим текстом) забезпечує управління всіма параметрами мовленнєвого сигналу і, таким чином, може генерувати текст за раніше невідомим текстом. Синтез реалізується шляхом моделювання та застосування аналогової або цифрової техніки. Так, синтез усного мовлення потребує імітації звуків мовлення в аналоговому або цифровому форматі, а генерація комп'ютером письмового тексту потребує достатнього розуміння принципів лінгвістичної організації тексту, узусу, контекстуального слововживання і под.

Перекодування ж тексту є аспектом комп'ютерної лінгвістики, що стосується безпосередньо перекладу, і тому становить особливий інтерес для перекладачів. Варіантів комп'ютерного перекладу тексту, як уже зазначалося, два – машинний та автоматизований.

Машинний переклад – процес перекладу текстів, поданих у письмовому або усному вигляді, з однієї природної мови іншою за допомогою спеціальної комп'ютерної програми. Існує чотири види машинного перекладу (рис. 4).



Рис. 4.

За умов постредагування вихідний текст повністю перекодується машиною, а людина-редактор корегує результати. За умов прередагування людина первинно пристосовує текст до обробки його машиною (прибирає можливі неоднозначності, спрощує та структурує текст), після чого починається програмка обробка. Інтерредагування передбачає, що людина втручається в процес перекладу, вирішуючи складні випадки та проблемні питання. Змішанні системи означають комбінування викладених вище варіантів.

Взагалі думки про те, що можна використовувати ЕОМ для вирішення питань перекладу була висловлена ще 1947 року. Ще за 7 років, 1954, стався так званий Джорджтаунський експеримент, відомий як перша публічна спроба демонстрації машинного перекладу. Інтелектуальна комп'ютерна система, що використовувалася тоді, містила словник з 250 слів, граматику з 6 правил і базу знань перекладу із кількох простих фраз. Незважаючи на її простоту, вона спричинила справжній бум у лінгвістичному світі і стимулювала похідні розробки у багатьох розвинених країнах. Вже у середині 1960х років американськими розробниками було представлено дві системи машинного перекладу з російської мови (зважаючи на тодішні актуальні

обставини холодної війни і потреби розвідки це було більш ніж логічним): MARC (розробка технічного департаменту ВВС) та GAT (розроблена вченими-дослідниками Джорджтаунського університету). Однак, з огляду на низьку якість перекладів, фінансування та стимуляція цих проектів були не надто щедрими, що пригальмувало розвиток даної галузі.

Якість машинного перекладу залежить від самого програмного забезпечення, але також і від тематики, стилістики, граматики, синтаксису та лексики тексту, спорідненості мов оригіналу та перекладу та інших факторів. Науково доведено практичним чином, що, хоча художній переклад внаслідок багатозначності тексту і контексту не піддається якісному машинному перекладу, деякі типи технічних або офіційно-ділових текстів отримали машинний переклад прийнятної якості, що потребував мінімального втручання редактора\коректора.

Серед відомих програм, використаних для машинного перекладу, слід назвати Trados, PROMT, Multitran, SmartCAT, Google Translate, DejaVu X3, а також українські перекладацькі системи Pragma, РУТА та російський Яндекс-Переводчик.

У будь-якому випадку ступінь залучення людини до процесу машинного перекладу мінімальний. Ситуація ж з автоматизованим перекладом інша.

Автоматизований переклад – переклад текстів людиною на комп'ютері із застосуванням комп'ютерних технологій. Від машинного перекладу він, власне, відрізняється тим, що процес перекладу повністю здійснено людиною, комп'ютер залучається лише до того, щоб пришвидшити процес або покращити його якість.

Цікаво, що ідея автоматизованого перекладу в сирому вигляді була висловлена ще 1933 року, коли радянський вчений П.П. Троянський виклав ідею розробки машини для підбору та друку слів при перекладі з однієї мови іншою. Пристрій складався з стола з похилою поверхнею, перед яким був закріплений фотоапарат, синхронізований із друкарською машинкою. На поверхні стола знаходилося «поле глосарію» – вільно рухома пластина з надрукованими словами трьома, чотирма і більше мовами. Зрозуміло, що на той момент ідея була технічно не реалізована.

На даний момент найрозповсюдженішими способами застосування інтелектуальних комп'ютерних систем для здійснення автоматизованого письмового перекладу є робота із словниками, глосаріями, пам'яттю перекладів та корпусами. Окрім того є спеціалізоване програмне забезпечення, що застосовується для вирішення конкретних завдань, наприклад, аудіовізуального перекладу, локалізації продуктів програмного забезпечення, редагування текстів.

Автоматизований переклад охоплює ціле поле інструментів, до його інструментарію належать:

- програми для перевірки правопису, які можуть бути вбудовані в текстові редактори, наявні на ПК, або існувати як окремо встановлене програмне забезпечення (Ursa Spell Checker, LanguageTool);
- програми для перевірки коректності граматичного оформлення тексту («Орфо»);

– програми для створення та навігації термінологічних баз в електронній формі, електронних таблиць Excel або баз даних FileMaker.

– електронні словники, тлумачні або багатомовні (ABBYU Lingvo, Multitran, Ectaco);

– термінологічні бази даних, що містяться на зовнішніх носіях або доступні онлайн, наприклад, TERMIUM;

– програмне забезпечення для керування проектами за перекладу, яке стане в пригоді менеджерам перекладачів та співробітникам перекладацьких агенцій, що стикаються з великим потоком запитів клієнтів (наприклад, Translation Office 3000).

Зупинимось детальніше на кожному із вищезазначених інструментів.

Призначення програм для перевірки правопису, синтаксису та пунктуації полягає в тому, що вони спрямовані на знаходження у тексті помилок орфографічного, граматичного, частково стилістичного та морфологічного характеру та їх виправлення. Сьогоднішні програми такого гатунку існують також у вигляді плаг-інів (розширень) для браузеру, що ним ви користуєтесь на домашньому ПК. Зазвичай у схему лінгвістичного аналізу таких програм входить розгляд граматики, стилю, пунктуації, великих та малих літер, типографіка, логічні помилки та, звичайно, перевірка орфографії.

Програми для створення та навігації термінологічними базами даних покликані структурувати словниковий запас під час створення перекладачем глосаріїв на ту чи іншу тематику. Наприклад, якщо перекладач має обробити текст з орнітології, йому буде доречно всі незнайомі слова з даної тематики винести до власноруч створеного **глосарію** – *міні-словника з певної спеціалізованої тематики*. Програмне забезпечення спеціального характеру покликане допомогти структурувати цей глосарій, упорядкувати лексику за алфавітом, винести до внутрішнього пошуковика та ін. Також вони можуть бути корисними для створення **тезаурисів** – *зводів даних, що в значному ступені повноти охоплюють поняття, визначення і терміни спеціальної області знань або сфери діяльності*. На відміну від словника, тезаурус допомагає

перекладачеві виявити зміст не тільки за допомогою розуміння слова, але й шляхом співвіднесення й зіставлення слова з іншими поняттями та їх групами. Тут і стає в пригоді перекладачеві комп'ютерна лінгвістика та III.

Електронних словників наразі існує дуже велика кількість, проте вони побудовані за двома основними принципами – 1) принцип введення слова за допомогою пристроїв вводу ПК (ABBYU Lingvo, Polyglossum, GoldenDict, Slovoed, Multitran) і 2) принцип спливаючого віконця при наведенні курсору миші разом із натисканням комбінації клавіш (iTranslate, StarDict, WordNet). Перший тип більш розповсюджений, до того ж, у найкращі словники першого типу вбудована також і функція словників другого типу. Так, зокрема, комбіновані функції має ABBYU Lingvo X+. Призначення електронного словника полягає у видачі у відповідь на запит користувача перекладу того чи іншого слова (або сталого словосполучення) із бази знань (словникової бази програми). При цьому словникові бази можуть бути завантажені у фізичну пам'ять комп'ютера, або ж програма може використовувати бази, розміщені онлайн (при цьому недоліком є необхідність постійного підключення ПК до Інтернету). Внутрішня структура електронного словника оформлена як база даних, що містить словникові статті, що дозволяє швидко знайти потрібне слово з урахуванням морфологічних особливостей, а також змінити напрям перекладу (у випадку, якщо це багатомовний словник). У принципі електронні словники є найчастіше використаним інструментом автоматизованого перекладу.

Нарешті, програмне забезпечення для керування проектами з перекладу стане в пригоді перекладачам, що одночасно беруться за кілька замовлень, або хочуть вести упорядковану клієнтську базу, бухгалтерію робочого процесу та контакт-листи. Програми такого роду спеціально структуровані для того, щоб містити різноманітні дані про клієнтів, проекти, строки їх виконання, дедлайни, історію перекладів, розцінки та розрахунки і под. Менеджери бюро перекладів використовують такі програми для відстеження професійної діяльності персоналу, контактів із замовниками і в широкому спектрі інших випадків.

Список літератури:

1. Андреев Н. Д. Статистико-комбинаторные методы в теоретическом и прикладном языковедении / Н. Д. Андреев. – Ленинград: Наука, 1967. – 408 с.
2. Арапов М. В. Математические методы в исторической лингвистике / М. В. Арапов, А. А. Херц. – Москва: Наука, 1974. – 192 с.
3. Большакова Е. И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: Учебное пособие / Е. И. Большакова, Е. С. Клышинский. – Москва: МИЭМ, 2011. – 272 с.
4. Волкова И. А. Введение в компьютерную лингвистику. Практические аспекты создания лингвистических процессоров / И. А. Волкова. – Москва: Изд-во Московского университета, 2006. – 45 с.
5. Жуков Д. А. Мы-переводчики / Д. А. Жуков. – Москва: «Знание», 1975. – 112 с.
6. Зубов А. В. Информационные технологии в лингвистике / А. В. Зубов, И. В. Зубова. – Москва: Издательский центр «Академия», 2004. – 208 с.
7. Карачун В. Я. Російсько-український словник з інформатики та обчислювальної техніки / В. Я. Карачун, Г. Г. Гульчук. – Київ: Видавництво «РОСЬ», 1994. – 363 с.
8. Комарова З. И. Методология, метод, методика и технология научных исследований в лингвистике / З. И. Комарова. – Екатеринбург: Изд-во УрФУ, 2012. – 818 с.
9. Марчук Ю. Н. Компьютерная лингвистика: Учебное пособие / Ю. Н. Марчук. – Москва: АСТ, 2007. – 317 с.
10. Панов Д. Ю. Автоматический перевод / Д. Ю. Панов. – Москва: Изд-во Академии Наук СССР, 1958. – 74 с.
11. Сова Л. З. Аналитическая лингвистика и типология / Л. З. Сова. – Санкт-Петербург: Изд-во Политехн. ун-та, 2007. – 378 с.

12. Шайкевич А. Я. Введение в лингвистику / Анатолий Янович Шайкевич. – Москва: Издательский центр «Академия», 2005. – 400 с.
13. Шемякин Ю. И. Начала компьютерной лингвистики / Ю. И. Шемякин. – Москва: Изд-во МГОУ, 1992. – 81 с.
14. Baker A. E. Linguistics / A. E. Baker, K. Hengveld. – Boston: Wiley-Blackwell, 2012. – 469 с.

Греков В.О.

Черноморский государственный университет имени Петра Могилы

РОЛЬ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ В ОБУЧЕНИИ И СОВЕРШЕНСТВОВАНИИ НАВЫКОВ ПЕРЕВОДА

Аннотация

Данная статья содержит анализ роли такой дисциплины, как компьютерная лингвистика, в обучении и работе профессионального переводчика. Предоставлены графические данные, наглядно иллюстрирующие материал, причины преподавания данной дисциплины в ВУЗах для студентов специальности «Перевод» и на факультетах лингвистической направленности, а также проанализировано на конкретных примерах то, как компьютерная лингвистика представляет интерес для начинающего и профессионального переводчика, и чем именно. Статья содержит толкование базовых понятий, а также конкретные примеры с пояснением необходимости изучения компьютерной лингвистики непосредственно для переводчиков. Исследован также механизм вовлечения языковедческих знаний в процесс перевода. Статья будет интересна как студентам филологических специальностей, так и специалистам в области перевода.

Ключевые слова: языкознание, компьютерная лингвистика, перевод, профессиональный переводчик, поисковик, алгоритм, корпус языка, распознавание речи, синтез речи.

Grekov V.O.

Petro Mohyla Black Sea National University

THE ROLE OF COMPUTATIONAL LINGUISTICS IN STUDYING AND ENHANCING TRANSLATION SKILLS

Summary

The article features an analysis of the role of Computational Linguistics as a scientific discipline in the preparation and operation of a professional translator and interpreter. It contains visuals enabling the understanding of the concepts given in the article, basic principles for lecturing the course at universities having «Translation and Interpreting» among their classes at foreign language departments. It also contains the analysis of basic notions of the discipline, as well as illustrative examples of various tools and the explanation of what interest they are for beginners and professional interpreters. The author also studied the mechanisms of outsourcing theoretical data in the process of translation. The article may be of interest both for the students of philology and those employed in the field of translation.

Keywords: linguistics, computational linguistics, interpreting, professional translator, search engines, body of language, speech recognition, speech synthesis.